THE ROLES OF BELIEF, EVIDENCE, PERSPECTIVE, AND INDIVIDUAL DIFFERENCES IN SCIENTIFIC EVALUATIONS.

A Thesis Submitted to the College of

Graduate Studies and Research

in Partial Fulfillment of the Requirements

for the Degree of Master of Arts

in the Department of Psychology

University of Saskatchewan

Saskatoon

By

Erin Leigh Beatty

PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

DISCLAIMER

Reference in this thesis/dissertation to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement, recommendation, or favoring by the University of Saskatchewan. The views and opinions of the author expressed herein do not state or reflect those of the University of Saskatchewan, and shall not be used for advertising or product endorsement purposes.

Requests for permission to copy or to make other uses of materials in this thesis in whole or part should be addressed to:

Head of the Department of Psychology
University of Saskatchewan
Saskatoon, Saskatchewan S7N 5A5
Canada

OR

Dean
College of Graduate Studies and Research
University of Saskatchewan
107 Administration Place
Saskatoon, Saskatchewan S7N 5A2
Canada

Abstract

Reasoners who adopt the perspective of another can increase the proportion of logically valid inferences they make (Thompson, Evans, & Handley, 2005). A possible explanation is that shifting perspective promotes analytic reasoning. If this were the case, then shifting perspectives should also reduce the belief-bias effect. Furthermore, strong evidence should be preferred over weak evidence. To test this, 256 participants read twenty-four research descriptions that varied in evidence quality and degree of personal belief content. Participants indicated whether the data supported the researcher's hypotheses. Belief bias was reduced when participants evaluated the data from the researcher's perspective relative to their own. Evidence strength was an important determining factor in decision-making and it was sensitive to perspective and individual differences.

Acknowledgements

I wish to gratefully acknowledge the support and guidance provided by my advisor Dr. Valerie A. Thompson and my committee members Dr. Jamie I.D. Campbell and Dr. Lorin J. Elias. I would like to thank Dr. Eric Neufeld for acting as my external examiner. I would like to thank Gregory P. Krätzig, Arron W. S. Metcalfe, and Jamie A. Prowse Turner for numerous stimulating discussions, help with stimuli development and general advice. I would like to recognize David Noete's valuable programming contribution without him the counterbalancing would not have been feasible. I would also like to acknowledge the Natural Sciences and Engineering Research Council (NSERC) for the financial support for this project. I wish to recognize the study participants for their time and contribution. Finally, I wish to thank my parents for their support and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

TASS	The Autonomous Set of Systems
AOT	Active Open-minded Thinking Scale
CRT	Cognitive Reflections test
SILS	Shipley Institute of Living Scale
EBI	Epistemological Belief Inventory
EBI-SK	Epistemological Belief Inventory Subscale simple knowledge
EBI-CK	Epistemological Belief Inventory Subscale certain knowledge
EBI-IA	Epistemological Belief Inventory Subscale innate ability
EBI-OA	Epistemological Belief Inventory Subscale omniscient authority
EBI-OA	Epistemological Belief Inventory Subscale quick learning
GDMS	General Decision Making Scale
GDMS-R	General Decision Making Scale Rational subscale
GDMS-I	General Decision Making Scale Intuitive subscale
GDMS-D	General Decision Making Scale Dependent subscale
GDMS-A	General Decision Making Scale Avoidant subscale
GDMS-S	General Decision Making Scale Spontaneous subscale
EUA	Epistemological Understanding Assessment
EUA-A	Epistemological Understanding Assessment Authoritative designate
EUA-E	Epistemological Understanding Assessment Evaluative designate
EUA-M	Epistemological Understanding Assessment Multiplist designate

1. INTRODUCTION

Our objective was to investigate how individuals overcome their propensity to make belief based decisions. Optimally reasoners should evaluate evidence based on the strength of that evidence and not their personal beliefs about the evidence. Belief bias is the tendency to evaluate conclusions based on prior beliefs independent of the evidence supporting the conclusion (Torrens, Thompson & Cramer, 1999). Belief bias is a widely replicated phenomenon that is difficult to eliminate (Newstead, Pollard, Evans, & Allen, 1992; Evans, Newstead, Allen, & Pollard, 1994). Past research that used instructions emphasizing logical necessity and the importance of ignoring personal beliefs was only marginally successful in reducing belief bias (Newstead et al., 1992; Evans et al., 1994) other research has found that reasoning performance was influenced by perspective change (Beatty & Thompson, 2007; Thompson, Evans & Handley, 2005). Individual differences, perspective taking and evidence quality were used in the current investigation of belief bias in scientific reasoning.

1.1 Dual Process Theory

The heuristic-analytic theory of reasoning is a generally acknowledged theoretical framework for explaining belief bias. Typically, the dual systems are described in the following way. Heuristic processing is quick, effortless, relies on gist based processing and often includes stereotypes. Analytic processing is slow, resource demanding and difficult to engage. Activities like hypothetical thinking require analytic processing. The output of heuristic processes is accepted as a default unless analytic processes intervene. Heuristic processes can produce belief bias by quickly producing examples that seem highly salient but are not normatively warranted (De Neys, Vartanian, & Goel, 2008).

Such is the case when individuals buy lottery tickets with the expectation of winning. Lottery winners are far the more salient outcome as they appear on television and the news media whereas the thousands of individuals that lose are not equally salient. Individuals fail to account for the base rate of lottery winners to losers when purchasing tickets. Furthermore, individuals feel their tickets are luckier than those belonging to others in the lottery (Risen & Gilovich, 2007). The latest versions of dual process theory make it clear that the heuristic and analytic systems are made up of multiple components. It is necessary to have a third type of process that determines which of the heuristic or analytic processes gets to control the response (Evans, 2009; Thompson, 2009; Stanovich, 2009).

Stanovich (2009) proposed a tri-process theory. Heuristic processes were renamed the autonomous set of systems (TASS) to represent that they were a heterogeneous set. TASS includes domain general products of implicit learning and conditioning as well as rules, stimulus discriminations and well practiced decision-making principles that have become automatic. TASS is autonomous in that it responds to stimuli and is not dependent upon input format nor is TASS controlled by the analytic processes. Despite the name change the form and function of TASS remains relatively unchanged, as it was the analytic processes where the greatest level of revision occurred. The analytic processes have been redefined as two levels of processing: The reflective level and the algorithmic level. The reflective level is responsible for the larger goal state, epistemic values, and critical thinking skills. The reflective mind is proposed to associate with preferences in thinking disposition. It is the reflective mind that is the new member of the dual process theory. Its role is to mediate the individual's goal state into an action that

results in either using a TASS output or committing a TASS override. The algorithmic level refers to a traditional conceptualization of analytic thinking. It is associated with measures of intelligence and cognitive capacity it is this level that is responsible for overriding the TASS outputs but it itself is subordinate to the reflective level.

In order to better understand the proposed reflective mind, and the hypothesis that it mediates analytic intervention, it is necessary to engender a conflict between the heuristic TASS output and the analytic (algorithmic level). One common method is to put beliefs in conflict with another dimension that is presumed to rely on analytic processes. In our case, it was the quality of the evidence participants were presented with. The measures of individual differences in thinking style, epistemology and decision making style provide an avenue of investigation into their proposed association with the reflective mind. The goal was to examine variables that affect the probability and quality of heuristic/TASS override. We made use of task manipulations that we proposed cued the need for analytic thinking, as well as individual differences in reflective thinking.

1.2 Belief Bias

Belief bias is easily replicated but its causes are not well understood. Individuals' prior beliefs impact reasoning by influencing these same individuals to accept conclusions consistent with their prior beliefs and reject belief inconsistent conclusions. A biased decision is one that is independent of the supporting evidence quality. The available research suggests that belief bias is not linked to the efficiency of algorithmic processing, suggesting that we need to look at reflective processing as a way to explain intervention (West, Toplak, & Stanovich, 2008).

Klaczynski and Narasimbam (1998) found that scientific evidence that supported personal beliefs was rated as valid and stronger compared to belief challenging evidence. Participants could objectively evaluate scientific evidence but unequally applied their skills to belief consistent and inconsistent material. Klaczynski and Lavallee (2005) found that belief consistent arguments were more persuasive and considered valid more often than belief threatening arguments. It is clear that beliefs have an impact on the perception and evaluation of evidence and arguments. These effects are not consistent across participants and any mediating factors that can account for this variability should be explored.

1.3 Individual Differences

Belief bias is not strictly determined by differences in cognitive ability (Klaczynski & Robinson, 2000; West, Toplak, & Stanovich, 2008). Klaczynski and Gordon's (1996) study of statistical reasoning in adolescents found that cognitive ability was not connected to biases in the adolescents' reasoning. Individuals were better at critically evaluating belief inconsistent evidence compared to their evaluations of belief consistent evidence.

A method for indexing reflective processes was required to explore the hypothesis that a predisposition to belief bias may be linked to these processes. As suggested by Stanovich (2009) measures of thinking dispositions is hypothesized to correlate with the reflective mind. Thinking style was defined here, as a tendency to think in a specific way (Klaczynski & Robinson, 2000). Thinking style must have some mechanism for how it impacts performance. It is best thought of as what an individual thinks is adequate to support their conclusion (Kuhn, 1991). Individuals differ in their level of motivation;

those who want to confirm their beliefs ought to respond differently than those who are open minded about material that conflicts with their beliefs. Thinking style measures can assess how people report they think and the value they place on changing one's opinions to reflect available evidence. Most thinking disposition measures use self-reports, so it is beneficial to use multiple measures to look for reporting consistency. Thinking dispositions are not distinct categorical entities rather they are an inclination toward one pattern at the expense of another. Stanovich and colleagues (1997, 1998, 2007) investigate thinking styles using the Actively Open Minded Thinking questionnaire (AOT). The AOT quantifies the value an individual reports they place on active thinking, self-reflection, belief change, and alternate explanations. Thinking disposition has predicted the magnitude of belief bias in formal (Stanovich & West, 1997, 1998) and informal reasoning tasks (Klaczynski & Robinson, 2000; Sa, Kelley, Ho, & Stanovich, 2005). We should observe a relationship between measures of thinking dispositions and belief bias such that those individuals who are active open-minded thinkers should show less belief bias than those that do not report that type of thinking preference. Furthermore, these active open-minded thinkers should also correlate positively with indices of analytic thinking. There are a variety of approaches to define and measure thinking dispositions. Prior work has focused on only one at a time, but our goal was to adopt a more integrative view. Below, a number of approaches are described from epistemology to cognitive patience.

Everyone has a way they think about knowledge this is their individual epistemology. An individual's epistemology can influence their reasoning performance (Kuhn, 1991). The evaluation of evidence is hypothesized to directly connect to an

individual's belief in the certainty of knowledge. Kuhn (1991) proposed three types of epistemological theories. Absolutist epistemologies view knowledge as objective and subsequently positions are either right or wrong. Multiplist epistemologies view knowledge as subjective; therefore, one person cannot be more certain than another. Evaluative epistemologies view knowledge as subjective and that it can be objectively evaluated based on the supporting evidence. Kuhn, Cheney and Weinstock (2000) assessed an individual's epistemological understanding across five dimensions: personal taste, aesthetics, values, truth about the social world, and truth about the physical world. Scraw, Dunkle and Bendixen (1995) also have a measure that assesses an individual's beliefs on five subscales: simple knowledge, certain knowledge, innate ability, omniscient authority and quick learning. We should observe a relationship between measures of epistemology and belief bias such that those individuals who have more complex epistemologies should show less belief bias than those that report less complex epistemological views. Moreover, those with complex epistemological views should also correlate positively with indices of analytic thinking.

The cognitive reflections test (CRT) by Frederick (2005) is thought to measure cognitive patience along with the general decision making scale (GDMS) by Scott and Bruce (1995) both provide avenues of investigation. We should observe a relationship between the performance-based measure of cognitive patience and belief bias such that those individuals who have more cognitive patience should show less belief bias than those who are not patient. Additionally, there should be a positive correlation between cognitive patience and indices of analytic thinking. The GDMS subscales may correlate with belief bias such that those individuals who have more impulsive decision making

style should show more belief bias than those that report more deliberate decision making styles. Those subscales may correlate with our indices of analytic thinking. These measures of epistemology, decision-making, and cognitive patience are novel inclusions that provide fresh areas of investigation of individual differences in reasoning. Thus, they may provide additional measures to index the reflective mind.

1.4 Scientific Reasoning and Perspective

Belief bias is a notoriously robust phenomenon that is difficult to eliminate, even by instruction (Newstead, Pollard, Evans, & Allen, 1992; Evans, Newstead, Allen, & Pollard, 1994). A small number of studies have found that manipulations of decision perspective have been effective in reducing belief bias (Greenhoot, Semb, Colombo & Schreiber, 2004; Beatty & Thompson, 2007; Thompson, Evans & Handley, 2005). It is hypothesized that the perspective manipulation moderated belief bias by increasing the amount of analytic processing the participant conducted.

Greenhoot, Semb, Colombo and Schreiber (2004) presented participants with the method and results of hypothetical experiments and then asked them to make judgments about belief-consistent or belief-inconsistent cause and effect relationships. In a within-subjects design participants were asked to make causal conclusions from two perspectives, their own and that of the researcher. The conclusion options were either consistent or inconsistent with the students' beliefs. Most students made accurate predictions about the belief consistent problem but less than a quarter of the participants made accurate predictions about the belief inconsistent problem. Participants were more accurate on the belief inconsistent problem from the researcher's perspective than when they made the conclusion from their own perspective. However, there were limitations to

this study. The belief condition was confounded with the study outcome. The belief consistent problem had evidence that supported a positive correlation whereas the belief inconsistent problem had evidence that supported a null relationship.

Research by Thompson, Evans and Handley (2005) found effects related to decision perspective comparable to Greenhoot et al.'s (2004) using a between subjects design. In a conditional argument task individuals accepted more inferences from the writer's perspective than from their own perspective. Thompson et al. interpreted this finding as individuals were more likely to accept the truth of the premise from the writer's perspective than from their own which could lead to analytic, decontexualized reasoning. Importantly, acceptance rates by inference validity was only significant for the group that responded from the writer's perspective and not for the group who responded from their own perspective. Reasoning from one's own perspective may promote contextualized processes that make use of past experience independent of the value that past experience has in the current context. Reasoning from another's perspective may reduce belief bias.

Beatty and Thompson (2007) used a scientific reasoning task that combined manipulations of belief and perspective. Individuals evaluated two belief-consistent and two belief-inconsistent vignettes. Each vignette described a research hypothesis, method and result. Participants made one conclusion from their own perspective and one from the researcher's perspective. Beatty and Thompson found that individuals were less biased when evaluating research vignettes from the researcher's perspective. They also found that high scores on the Actively Open Minded Thinking scale were correlated with the ability to identify flaws in the evidence, but not with belief bias. While this study

disentangled belief from the evidence conclusion relationship present in Greenhoot et al.'s (2004), it too had its limitations. Participants were very likely to choose a conclusion that supported the hypothesis in the vignette this was a disconcerting result, given the poor quality evidence presented in all conditions. Therefore, the reduction in belief bias might be a bias to accept the available conclusion independent of personal beliefs rather than a reduction in bias due to increase analytic processing.

The results of Thompson et al. provide supporting evidence that changing decision perspective increases analytic processing as inference validity was only significant from the writers perspective in their study. Given that the scientific evaluation task of Beatty and Thompson (2007) was substantively different from conditional arguments, a performance comparison that includes an evidence quality manipulation would improve the specificity of the conclusion that can be drawn from the results.

1.5 The Current Study

The purpose of the current study was to investigate variables that underlie propensity to override a heuristic TASS response and intervene with analytic response in a scientific evaluation task. Individual differences in thinking style, epistemology and decision making style were used to explore their relationship with belief bias in a scientific reasoning task. The proposed mechanism for the interface between thinking style and reasoning performance was the proficiency that individuals engage in analytic processing. Specifically, that analytic processing was more likely to result in accurate evaluations of reasoning stimuli as compared to heuristic processing.

In the current experiment participants evaluated 24 experiment scenarios that had a range of possible degrees of prior beliefs. Stimuli were developed expressly so that

some items would be highly belief consistent or belief inconsistent while others would be neutral. As with Beatty and Thompson's previous work participants made two conclusions, one from the perspective of the researcher and one from their own perspective. Furthermore, individual differences in AOT, CRT, GDMS, and two measures of epistemological understanding provided indices of the reflective mind, and algorithmic processing was indexed by the Shipley Institute of Living Scale (SILS) (Zachary, 1991) as proposed in Stanovich (2009). Based on the previous literature review there are multiple predictions being made.

Hypothesis one was that the perspective manipulation increases the amount of analytic processing the participant conducts thereby moderating belief bias. There should be less belief bias and greater reliance on evidence quality from the researcher's as opposed to the participant's perspective. Additionally, experiments supported by strong evidence should be endorsed at higher rates than those experiments supported by weak evidence.

Hypothesis two was that the epistemological measures index the reflective mind and therefore, should be associated with the predisposition to engage analytic thinking. Thus, they should be correlated with both belief bias and reliance on evidence quality.

Hypothesis three was that if the individual differences tests measure similar constructs then they should correlate with each other and with similar performance dependent variables. Conversely, if these different measures are not correlated there are two possible interpretations, either they represent independent constructs or they do not represent characteristics that have a relationship with the reasoning process.

2. Method

2.1 Participants

There were two groups of participants. One group was recruited from the University of Saskatchewan community including staff and students using posters and website advertisements. Participants were compensated at a rate of $5.00 per half hour of participation. There were 77 participants (34 men and 42 women) with a mean age of 22.13 years (*SD* = 5.25). Participants had an average of 15.24 years (*SD* = 2.30) of formal education. We asked participants to report the highest level of education achieved: 26.3% had completed high school, 1.3% had technical training beyond high school, and 51.3% had college or some university, 18.4% had an undergraduate degree, and 2.6% had a graduate degree. Seventy-six participants indicated English as their preferred language and one participant did not report language preference.

The second group of participants consisted of 179 individuals (35 men and 143 women, one did not report) with a mean age of 29.05 years (*SD* = 11.65). These participants were recruited online through advertisements on research sites and online classified sites. Participants had an average of 16.35 years (*SD* = 3.00) of formal education. We asked participants to report the highest level of education achieved: 0.6% had completed some high school, 9.0% had completed high school, 13.9% had technical training beyond high school, 37.1% had college or some university, 31.3% had an undergraduate degree, and 18.0% had a graduate degree. One hundred and seventy-five participants indicated English as their preferred language, one reported German English bilingualism, one reported Spanish and one did not report. Those participants who completed the study online did the exact same task as those who completed it in the lab with one exception; they did not complete the Shipley Institute of Living scale as the

cognitive ability measure was not licensed for online use nor was it feasible to do so due to its administration procedure.

2.2 Materials

2.2.1 Experiment evaluation task. The experiment evaluation task consisted of three parts; which will be described individually. The first part was the pretest, which was followed by the evaluation task and then the experiment evaluation short answer questions.

2.2.2 Pretest. The purpose of the pretest was to assess the participant's personal beliefs on the topics they would later evaluate in the experiment evaluation task. Participants indicated their agreement or disagreement with 24 statements on a seven point scale: 1 (Strongly Disagree) 2 (Disagree) 3 (Disagree Slightly) 4 (Neither Agree nor Disagree) 5 (Agree Slightly) 6 (Agree) 7 (Strongly Agree). The items referred to social issues such as criminal behavior and child education (e.g., students who enjoy school are more likely to succeed academically than students who do not enjoy school). Four items were taken from Stanovich and West (1997), two from Greenhoot et al. (2004) and the remaining eighteen were developed for the current experiment. These items represent the hypotheses for the experiments participants later evaluated.

2.2.3 Experiment evaluation. Participants evaluated twenty-four experimental scenarios. Each experiment scenario contained: a description of the researcher, a hypothesis, a piece of evidence, and four conclusions. For example:

> Dr. Kennedy, a psychologist, wanted to investigate school success. He used a questionnaire to evaluate students' academic experiences in high school. Dr. Kennedy hypothesized that students who enjoyed school would be more likely to succeed academically. Dr. Kennedy found that in Saskatoon, 90% of those students who enjoyed school graduated, whereas national statistics show that only 10% of those who did not enjoy school graduated.

What should **Dr. Kennedy** conclude based on the evidence provided here?
A) High school students who enjoy school are more successful than those who do not enjoy school.
B) High school students who enjoy school are less successful than those who do not enjoy school.
C) There is no relationship between high school students' enjoyment of school and their school success.
D) No conclusions are warranted.

What would **you** conclude based on the evidence provided here?
A) High school students who enjoy school are more successful than those who do not enjoy school.
B) High school students who enjoy school are less successful than those who do not enjoy school.
C) There is no relationship between high school students' enjoyment of school and their school success.
D) No conclusions are warranted.

Participants evaluated each of the four conclusions from their own and the researcher's perspective. The perspective that the evaluation was made from was randomly assigned. These four conclusions described a different possible relationship between the two variables described in the vignette. In half the scenarios, the researcher was of high authority and was low for the other half. The assignment of evidence type to authority and evidence quality conditions was randomly assigned.

Half of the experiments presented poor quality evidence and half better quality evidence. Four types of evidence were presented to increase the perceived validity of the experiments and therefore retain the participant's experimental interest. Each type had a strong and weak version. Statistical evidence described the degree of correlation between two variables. In the weak version of this evidence only one direction of the relationship was described whereas the strong evidence version described both directions of the relationship (e.g. Strong: statistics from the current study show that when the number of students that enjoyed school increased, graduation rates increased; conversely, when the

number of students that enjoyed school decreased, so did the graduation rates. Weak: statistics from the current study show that when the number of students that enjoyed school increased, graduation rates increased.)

A second type of evidence included information about the co-occurrence of cause and effect with either a large sample ($n > 96$) and thereby strong piece of evidence or a small sample ($n < 8$) thereby weak piece of evidence (e.g., Strong: of the 110 people who were sampled, all 55 of the students that enjoyed school graduated, but 0 of the 55 students that did not enjoy school graduated). A third type of evidence provided information about the co-occurrence of cause and effect. The strong version of the evidence included the converse probability whereas the weak version confounded the comparison with two different samples (e.g., Strong: 90% of those students who enjoyed school in Saskatoon graduated, whereas only 10% of those students in Saskatoon who did not enjoy school graduated. Weak: in Saskatoon, 90% of those students who enjoyed school graduated, whereas national statistics show that only 10% of those who did not enjoy school graduated). The final evidence type was a survey in which a large group from the population was surveyed. In the strong evidence version the group sampled was from the population the conclusions were generalized to and the evidence contained no explicit confounds. In the weak version of this evidence type the sample was restricted to a small subset of the total population to which the conclusion was generalized and it was confounded by geography, socioeconomic status etc. (e.g., Strong: In a province wide survey, those students that enjoy school graduated at higher rates than students that did not enjoy school. Weak: In a province wide survey of grade nine students, those students

that enjoyed school graduated from large schools at higher rates than students from small schools that did not enjoy school).

Participants chose one of four conclusion options for each experiment scenario. The first asserted a positive relationship between two variables, the second a negative relationship, the third a null relationship and the fourth asserted that no conclusions were warranted. The first option supported the researcher's hypothesis, whereas options two and three contradicted the hypothesis. The first was considered correct for the scenarios with strong evidence presented and the fourth was considered correct for the scenarios that had weak evidence.

2.2.4 Experiment evaluation short answer questions. Participants were asked a series of short questions requiring either a 'yes' or 'no' response or a short written response. The questions were related to the authority, evidence strength and perspective manipulations. For example; participants were asked, "Did you notice anything about the evidence quality?" (Yes/No) and "Did the evidence provided effect your decisions?" (Yes/No). Several other questions pertaining to authority and perspective were included. The purpose of these questions was to get some self report data as to what the participants were doing during the task.

2.3 Individual Differences Measures

2.3.1 Actively open-minded thinking scale (AOT). The AOT measures self reported tendency to engage active, flexible, hypothetical thinking (Stanovich & West, 2007; Stanovich & West, 1997; Sá, Stanovich & West, 1999). Participants indicated their agreement on a six-point scale on items such as "No one can talk me out of something I know is right" and "If I think longer about a problem I will be more likely to solve it."

thirty of the items are reverse scored. Scores were obtained by totaling the responses to all 41 statements.

2.3.2 Cognitive reflections test (CRT). The CRT is a three-item measure that is thought to assess cognitive patience (Frederick, 2005). Each question prompts a heuristically motivated response; however, this answer is incorrect. In order to be successful an individual must suppress that initial heuristically produced response and derive a response based on further analyses. One example item is: A bat and a ball cost $1.10. The bat costs $1.00 more than the ball. How much does the ball cost? ___ cents. Usually the initial response is that the ball must cost 10 cents, but that is incorrect.

2.3.3 Epistemological belief inventory (EBI). The EBI is a 28 item scale assessing an individual's beliefs on five subscales: simple knowledge (seven items; e.g. Most things worth knowing are easy to understand), certain knowledge (five items; e.g. What is true is a matter of opinion), innate ability (six items; e.g. Really smart students don't have to work as hard to do well in school), omniscient authority (five items; e.g. People shouldn't question authority) and quick learning (five items; e.g. Working on a problem with no quick solution is a waste of time). Five items were reverse scored. Participants indicated their level of agreement with a series of statements on a six-point scale from Strongly Disagree (1), Disagree (2), Slightly Disagree (3), Slightly Agree (4), Agree (5), Strongly Agree (6). Scores on each of the subscales were obtained by calculating the mean score on each of the subscales. High scores indicate naïve views, like a belief in certain knowledge and low scores indicate individuals who reject such a position (Schraw, Dunkle, & Bendixen, 1995).

2.3.4 General decision making style (GDMS). The GDMS scale consists of 25 statements that assess how individuals make important decisions. Participants are asked to indicated the degree to which they agreed or disagreed with each statement on a five point scale: Strongly Disagree (1), Somewhat Disagree (2), Neither Agree Nor Disagree (3), Somewhat Agree (4), Strongly Agree (5). The questionnaire has five items for each of the five subscales. Rational style is characterized by a search for logical alternatives and evaluations (e.g. I plan my important decisions carefully). Intuitive style is one that uses feelings and hunches (e.g. When making decisions, I rely upon my instincts). Dependant style relies on advice and direction from others (e.g. I often need the assistance of other people when making important decisions). Avoidant style is one that tries to evade the decision (e.g. I postpone decision making whenever possible). Finally, the spontaneous style attempts to complete the decision as quickly as possible (e.g. I generally make snap decisions). Scores were obtained by calculating the mean response on each of the subscales and an overall style was calculated by determining the highest score of each of the subscales (Scott & Bruce, 1995).

2.3.5 Epistemological understanding assessment (EUA). The EUA assess an individual's beliefs about knowledge and knowing. Participants are classified into three groups based on their responses to four sets of questions. For example:

Robin says the stew is spicy.
Chris says the stew is not spicy at all.

Can only one of their views be right, or could both have some rightness?
Only one right Both could have some rightness

If both could be right:
Could one view be better or more right than the other?
One could be more right than the other
One could not be more right that the other

The three groups are: Absolutist, multiplist, and evaluative (Kuhn, Cheney & Weinstock, 2000). The response “only one right” indicates an absolutist thinking style. The response “One could not be more right that the other” implies a multiplist thinking style. Evaluative thinkers endorse “both could have some rightness” and that “one could be more right than the other.” Participants were evaluated on five dimensions: Judgments of personal taste (Robin says warm summer days are nicest. Chris says cool autumn days are nicest.), aesthetic judgments (Robin thinks the first painting they look at is better. Chris thinks the second painting they look at is better.), value judgments (Robin thinks lying is wrong. Chris thinks lying is permissible in certain situations.), judgments of truth about the social world (Robin has one view of why criminals keep going back to crime. Chris has a different view of why criminals keep going back to crime.), judgments of truth about the physical world (Robin believes one book's explanation of what atoms are made up of. Chris believes another book's explanation of what atoms are made up of.).

As Kuhn, Cheney and Weinstock (2002) specified, participants were assigned to the categories of absolutist, multiplist, or evaluative for a specific judgment type when they answered a minimum of two of the three items the same. For example, if a participant endorsed the “only one right” for two of the three items in the values judgment dimension they would be classified as an absolutist on that dimension. The total number of times across all dimensions an individual gave absolutist, multiplist or evalulativist answer was added and a total value was obtained.

2.3.6 Shipley institute of living scale (SILS). The SILS is self-administered and consists of two subtests: The vocabulary subtest and the abstraction subtest (Zachary, 1991). The vocabulary subtest consists of 40 multiple-choice questions in which the

individual is asked to choose which of four words is most similar in meaning to a target word. Administration time for the subtest is 10 minutes. The abstraction subtest consists of 20 questions in which sequences of numbers, letters, or words with the final element in each sequence omitted. The individual is required to complete each of the sequences. Administration time for each of the subtests is 10 minutes and the SILS provides a valid alternative to more intensive IQ measurements. It has been highly correlated with the Wechsler Adult Intelligence Scale–Revised, Wechsler Adult Intelligence Scale III, and the Kaufman Brief Intelligence Test. (Bowers & Pantel, 1985; Zachary, Crumpton & Speigal, 1985; Weiss & Schell, 1991). The SILS exhibits good internal consistency, test-retest validity, and criterion-related validity (Bowers & Pantel, 1985; Zachary, Crumpton & Speigal, 1985; Weiss & Schell, 1991).

2.4 Procedure

Testing occurred in one session in the following order. Participants completed the Shipley Institute of Living Scale, demographics, pretest, experiment evaluation task, experiment evaluation short answer questions, Actively open minded thinking scale, Cognitive reflections test, Epistemological Beliefs Inventory, Epistemological Understanding Assessment, and the General Decision Making Style questionnaire. Research scenarios were randomly ordered. Participants were tested individually and in small groups. The experiment took approximately sixty minutes to complete.

Those participants who completed the study online completed the same measures with the exception that they did not complete the Shipley Institute of Living Scale. When the participants viewed an advertisement for the study on the internet, they were directed to the study webpage where they read and completed the standard consent procedures.

3. Results

The design of this experiment was quite complex involving many manipulated factors as well as measures of individual differences. As such, the analyses were broken into three sections. First, an ANOVA of scientific evaluation task was conducted. Second, a regression analysis of belief in the scientific evaluation task is reported. Belief was not included in the ANOVA as it was measured continuously. Finally, the analyses of individual differences and their relationship with performance on the scientific evaluation task were conducted.

3.1 Experiment Evaluation Task ANOVA

The factors of perspective, authority, and evidence quality were entered into a 2x2x2 repeated measures ANOVA. The dependant variable was conclusion acceptance, meaning the rate that participants endorsed the conclusion option that supported the researcher's hypothesis. There was a main effect of perspective such that participants were more likely to endorse a conclusion that supports the researcher's hypothesis from the researcher's perspective (66.0%) compared to their own perspective (60.6%), $F(1, 255)=31.75$, $MSE=.05$, $p<.001$. There was a main effect of evidence strength such that strong evidence (70.7%) was supported more often than weak evidence (55.9%), $F(1, 255)=139.20$, $MSE=.08$, $p<.001$. This result supported the Hypothesis One, which was that strong evidence would be endorsed at higher rates than weak evidence. There was a perspective by evidence strength interaction, $F(1, 255)=5.36$, $MSE=.01$, $p<.05$. The difference in acceptance rates between strong and weak evidence is greater from the researcher's perspective than the participant's perspective (See Figure 1). The perspective by evidence strength interaction supported the previous interpretation by Beatty and

Thompson (2007) and our current contention that asking participants to reason from another's perspective promotes analytic processes and greater reliance on evidence quality in the researcher's as opposed to the participant's perspective. The increased level of discrimination between evidence qualities from the researcher's perspective could be due to increased use of those analytic processes. There was no main effect of authority; participants were equally likely to support the researcher's hypothesis if the researcher was described as a psychologist (63.1%) or if they were described as a student (63.6%), $F(1, 255)=0.19$, $MSE=.07$, $p>.05$. No other interactions were significant.

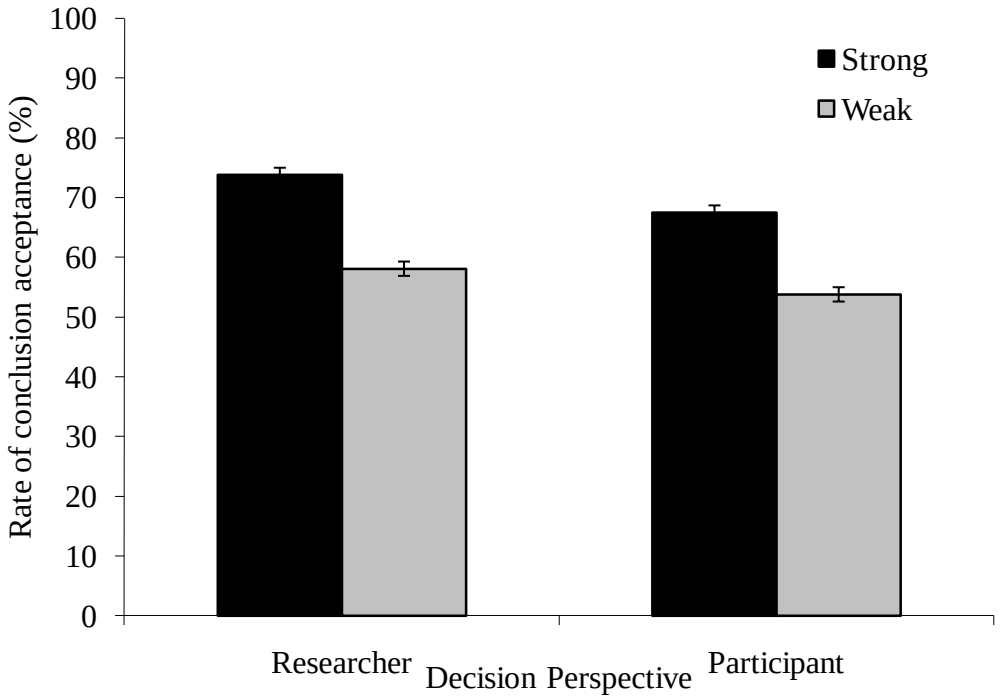

Figure 1. Perspective by evidence strength ANOVA interaction. Jarmasz and Hollands (2009) confidence intervals are reported.

3.2 Regression Analyses of Belief

The dimension of belief was assessed continuously by participant, as a result of this method of measurement belief could not be included in the ANOVA analyses and

was best evaluated within a regression model. We wanted to independently assess the effects of belief and the anticipated belief by perspective interaction for each participant's decisions; therefore, in accordance with Stanovich and West's (1997) analysis strategy two regressions were conducted for each participant for a total of 502 individual regressions. Two regressions were conducted for each person: One for the responses from the researcher's perspective and one for the responses from the participants' perspective. These were conducted in this manner so the effect of personal belief could be isolated by decision perspective. The independent variables of pretest belief ratings, authority level and evidence quality were entered into the regression equation simultaneously. Therefore, the regression coefficients for each independent variable were calculated controlling for all the other independent variables. The dependent variable for each regression was conclusion acceptance rates from each perspective. These analyses resulted in two beta weights for belief from the participant's and researcher's perspective. This belief beta weight was indicative of the degree of reliance on personal beliefs independent of evidence quality and experimenter authority. The mean belief beta weight from the researcher's perspective was .043 (*SD* =.213) and the mean beta weight from the participant's perspective was .095. (*SD* =.242). Although both of these values are small they were both significantly different from zero, $t(235) = 3.08$, SE=.01, $p<.01$; $t(240) = 6.07$, SE=.02, $p<.001$ respectively. Furthermore, they are significantly different from each other ($t(232) = -4.23$, $p<.001$). There was a great deal of variability in the regression beta weight values for individual's reliance on belief. These results nonetheless support hypothesis one that the perspective manipulation increases the amount of analytic processing the participant conducts thereby moderating belief bias in the conclusions.

3.3 Experiment Evaluation Short Answer Questions

We asked participants a series of questions about the experiment evaluation task after they had completed the 24 problems. When participants were asked: Did you notice that some of the experimenter's were doctors and others were students or journalists? (Yes/No) 86.5% of participants indicated 'yes' they had noticed. When participants were asked: Did that matter to you when you were making your decisions? (Yes/No) 94.7% of participants said no it did not matter. This last response pattern is consistent with the ANOVA analyses as variable of authority was not significant. When participants were asked: Did you notice anything about the evidence quality? (Yes/No) 86.1% of participants indicated 'yes' they had noticed. When participants were asked: Did the evidence provided affect your decisions? (Yes/No) 82.6% of participants 'yes' it did impact their decisions. This is consistent with the ANOVA findings of a main effect for evidence strength. To further investigate the perspective manipulation we then asked participants two general questions about their decision making process: When asked to decide from the experimenter's perspective what did you do? and when asked to decide from your perspective what did you do? These two questions were intended to get a general idea of what the participants could report about their performance. The participants written responses were coded to evaluate participants' use of the presented evidence. Responses were assigned a score ranging from 0 to 4. A response was given the code 4 if it contained reference to testing the researcher's hypothesis, evaluating the evidence, looking at the statistics, or logic. Code 3 was given if the response contained reference to any of the qualities of code 4 but also mentioned the use of personal beliefs. A response was given the code 2 if it only mentioned using personal beliefs, personal

background knowledge, opinions and gut feelings. Code 1 was used in a particular case where participants indicated that they pretended they were the researcher for example "put myself in their situation" or "used my own perspective as if I was the experimenter". Zero was coded if a response did not fit into these categories (e.g., incomplete statements, unintelligible responses, etc.). See Figure 2 for the complete coding results. Participants were more likely to report an evidence based (code 4) response for the researcher's perspective (70.0%) compared to their own perspective (58.9%), $t(206) = 2.77$, $p = .006$. Participants were more likely to report using some form of personal beliefs (code 3) from their perspective (13.4%) compared to the researcher's perspective (0%), $t(206) = -5.68$, $p < .001$. Additionally, they were more likely to report only using personal beliefs (code 2) from their own perspective (12.0%) compared to the researcher's perspective (2.4%), $t(206) = -4.07$, $p < .001$. Overall, participants reported more evidence base response patterns from the researcher's perspective and they reported more instances of using personal belief from their perspective. These results mirror the regression findings of greater belief effects from the participant's perspective and they also mirror the evidence by perspective interaction from the ANOVA.

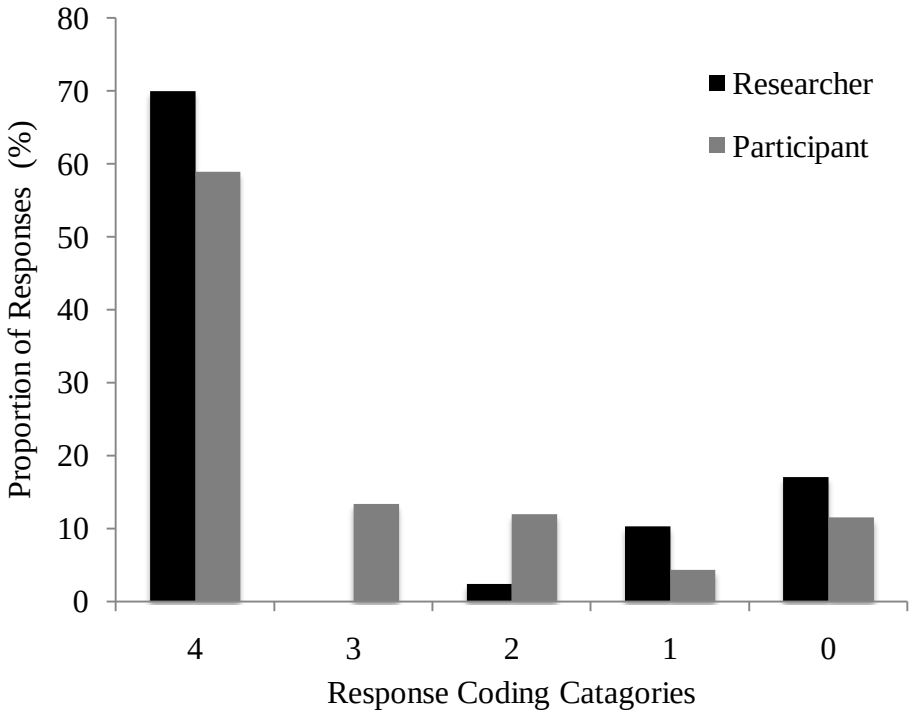

Figure 2. Experiment evaluation written response coding results.

Individual Differences

Participants completed a number of individual differences measures. The mean scores of these measures can be found in Table 1. These measures of individual difference were correlated with the standardized beta weights for belief, authority and evidence from each perspective (See Table 2). This analysis strategy was implemented to determine if any of these individual differences was related to the use of personal beliefs, reliance on authority, and sensitivity to evidence quality. These analyses directly relate to Hypothesis Two, that the epistemological measures index the reflective mind and therefore, should be associated with the predisposition to engage analytic thinking. Thus, they should be correlated with both belief bias and reliance on evidence quality.

Table 1
Individual difference measure means.

Measure	Mean	*SD*	Cronbach's Alpha
AOT	181.56	20.01	0.87
CRT	50.19	38.62	0.71
SILS IQ Estimate	114.34	5.62	-
EBI-SK	2.83	0.66	-
EBI-CK	2.69	0.81	-
EBI-IA	3.46	0.78	-
EBI-OA	3.29	0.75	-
EBI-QL	2.17	0.59	-
GDMS-R	4.06	0.53	-
GDMS-I	3.58	0.76	-
GDMS-D	3.39	0.81	-
GDMS-A	2.72	1.04	-
GDMS-S	2.71	0.84	-
EUA-A	1.23	1.67	-
EUA-E	7.72	4.27	-
EUA-M	5.84	4.18	-

Active open minded thinking scale (AOT), Cognitive Reflections test (CRT), Shipley Institute of Living Scale (SILS) IQ estimate, Epistemological Belief Inventory subscale simple knowledge mean (EBI-SK), subscale certain knowledge (EBI-CK), subscale innate ability (EBI-IA), subscale omniscient authority (EBI-OA), and subscale quick learning (EBI-OA), General decision making scale: rational (GDMS-R), intuitive (GDMS-I), dependent (GDMS-D), avoidant (GDMS-A), and spontaneous (GDMS-S), Epistemological Understanding Assessment: authoritative (EUA-A), evaluative (EUA-E), multiplist (EUA-M).

Table 2
Correlations between regression beta weights and individual differences.

	Researcher Perspective		Participant Perspective	
	Belief Beta Weight	Evidence Beta Weight	Belief Beta Weight	Evidence Beta Weight
SILS IQ estimate	-.138	.209+	-.027	.255*
AOT	-.102+	.150*	-.100	.117+
CRT	-.007	.176**	-.084	.221**
EBI-SK	.063+	-.177**	.087	-.132*
EBI-CK	.014	.019	-.005	-.044
EBI-IA	.085	.036	.004	.059
EBI-OA	-.005	-.119+	.034	-.122+
EBI-QL	.047	-.075	-.039	-.038
GDMS-R	-.017	.013	-.032	-.017
GDMS-I	.087	-.087	.102	-.150*
GDMS-D	.143*	.006	.117+	.007
GDMS-A	.176**	-.016	.079	.036
GDMS-S	.035	-.193**	-.023	-.167*
EUA-A	-.022	-.060	.007	.045
EUA-E	.041	-.083	.039	-.155*
EUA-M	-.060	.101	-.065	.144*

**Correlation is significant at the 0.05 level (2-tailed). **Correlation is significant at the 0.01 level (2-tailed). +Correlation is significant at the 0.05 level (1-tailed). Active open minded thinking scale (AOT), Cognitive Reflections test (CRT), Shipley Institute of Living Scale (SILS) IQ estimate, Epistemological Belief Inventory subscale simple knowledge mean (EBI-SK), subscale certain knowledge (EBI-CK), subscale innate ability (EBI-IA), subscale omniscient authority (EBI-OA), and subscale quick learning (EBI-OA), General decision making scale: rational (GDMS-R), intuitive (GDMS-I), dependent (GDMS-D), avoidant (GDMS-A), and spontaneous (GDMS-S), Epistemological Understanding Assessment: authoritative (EUA-A), evaluative (EUA-E), multiplist (EUA-M).*

The belief beta weight from the researcher's perspective was positively correlated with EBI simple knowledge subscale, GDMS dependant subscale, GDMS avoidant subscale, and negatively correlated with AOT. The belief beta weight from the participant's perspective was positively correlated with GDMS dependant subscale and with the researcher's perspective belief beta weight.

The evidence beta weight from the researcher's perspective was positively correlated with SILS IQ estimate, AOT, CRT, and negatively correlated with EBI simple knowledge subscale, EBI omniscient authority subscale, and GDMS spontaneous subscale. The evidence beta weight from the participant's perspective was positively correlated with SILS IQ estimate, AOT, CRT, EUA multiplist and with the researcher's perspective evidence beta weight. It was negatively correlated with EBI simple knowledge subscale, EBI omniscient authority subscale, GDMS intuitive subscale, GDMS spontaneous subscale, EUA evaluative.

The individual difference measures did correlate with the beta weights for belief although far more measures correlated with the beta weights for evidence. Those measures associated with the evidence beta weight displayed more consistent relationships from both perspectives compared to the belief beta weights. EBI simple knowledge subscale positively correlated with belief beta weight from the researcher's perspective and negatively with evidence from both perspectives.

Those participants with high AOT scores should be more likely to use analytic processes and distinguish between strong and weak evidence. The AOT was positively correlated with the evidence beta weight from the researcher's perspective meaning that high AOT individuals did use evidence more often than low scores, furthermore; it was

negatively correlated with belief beta weights from the participant's perspective. Low AOT individuals were more likely to use beliefs from their perspective compared to the high AOT scorers.

Hypothesis three was that the individual differences tests measuring similar constructs should correlate with similar performance dependent variables (See Table 3). The general pattern of the individual difference scores was that they tended to correlate with the evidence beta weights most frequently. The AOT and the GDMS were the only measures that correlated both with belief and evidence beta weights. The EBI did not correlate very consistently, with only two subscales reaching significance. The EUA was rather incoherent, the two scores that did reach significance did not make theoretical sense as the evaluative epistemology was negatively related to using evidence from the participants perspective and we expected the opposite finding. Also, the multiplist epistemology should have been related to the use of belief-based reasoning and instead was positively related to using evidence. The measures seem to be related to the use of analytic processes given that the stimuli was highly contextualized within belief laden content they seem to index the reflective level (Stanovich, 2009).

Looking next at the individual differences measures and the inter-scale correlations. The SILS IQ estimate was related to the AOT and CRT. The real standout scale from these findings was the AOT that was correlated with twelve scales and subscales. The EBI's subscales were all inter-correlated. While the EBI GDMS and EUA were all included to offer some diversity to the measures of individual difference they did not tend to offer anything beyond the AOT in associations with reasoning performance.

Table 3 *Inter-test correlations between individual difference measures and regression beta weights.*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1. SILS IQ estimate	-																					
2. AOT	.255*	-																				
3. CRT	.497**	.247**	-																			
4. EBI-SK	-.205	-.597**	-.143*	-																		
5. EBI-CK	-.170	-.378**	-.066	.187**	-																	
6 EBI-IA	-.022	-.177**	.122+	.137*	.064	-																
7. EBI-OA	-.163	-.377**	-.201**	.315**	.105	.032	-															
8. EBI-QL	-.067	-.530**	-.138*	.439**	.122	.303**	.089	-														
9. GDMS-R	-.019	.231**	.046	-.142*	-.043	.036	.109+	-.250**	-													
10. GDMS-I	-.188	-.136*	-.183**	.110+	-.063	.129*	.109+	.049	-.053	-												
11. GDMS-D	-.201	-.056	-.033	.063	.031	.067	.130*	-.032	.083	.060	-											
12. GDMS-A	-.224+	-.169**	-.007	.142*	.071	.158*	-.011	.268**	-.199**	.065	.324**	-										
13. GDMS-S	-.149	-.189**	-.091	.164*	-.031	.170**	-.056	.322**	-.380**	.391**	.076	.236**	-									
14. EUA-A	.038	-.180**	-.088	.010	.148*	.188**	-.017	.162*	-.029	-.018	-.036	.072	.047	-								
15. EUA-E	-.015	.092	-.064	.065	-.194**	-.208**	.065	-.055	.097	.142*	-.011	-.095	-.021	-.217**	-							
16. EUA-M	.069	.007	.104	-.091	.102	.154*	-.072	.019	-.090	-.117+	.040	.044	-.006	-.138*	-.876**	-						
17. Belief beta weight researcher	-.138	-.102+	-.007	.063+	.014	.085	-.005	.047	-.017	.087	.143*	.176**	.035	-.022	.041	-.060	-					
18. Authority beta weight researcher	-.028	.099	.042	-.092	.045	.045	-.107	-.013	.029	.039	-.029	-.018	.033	-.040	-.030	.041	-.119+	-				
19. Evidence beta weight researcher	.209+	.150*	.176**	-.177**	.019	.036	-.119+	-.075	.013	-.087	.006	-.016	-.193**	-.060	-.083	.101	.000	-.154*	-			
20. Belief beta weight participant	-.027	-.100	-.084	.087	-.005	.004	.034	-.039	-.032	.102	.117+	.079	-.023	.007	.039	-.065	.719**	-.106	-.054	-		
21. Authority beta weight participant	-.045	.014	-.004	-.057	.073	.017	-.014	-.043	.079	-.004	.002	-.040	-.006	-.062	-.049	.064	-.109	.804**	-.153*	-.112	-	
22. Evidence Beta weight participant	.255*	.117+	.221**	-.132*	-.044	.059	-.122+	-.038	-.017	-.150*	.007	.036	-.167*	.045	-.155*	.144*	-.034	-.142*	.691**	-.085	-.100	-

*Correlation is significant at the 0.05 level (2-tailed). **Correlation is significant at the 0.01 level (2-tailed). +Correlation is significant at the 0.05 level (1-tailed).

4. Discussion

The results of this study confirm that a relatively simple manipulation of perspective increases an individual's propensity to engage in analytic processing. This is evidenced by the interaction of perspective and evidence strength, which indicated that evidence strength had the greatest impact from the researcher's perspective. The effects of belief were greater from the participants perspective compared to the researcher's perspective. These findings extend those of Beatty and Thompson (2007) and Thompson, Evans and Handley (2005). It provides credible evidence from an entirely different task that supports the idea proposed by Thompson et al. that participants are more likely to accept the truth of the premise when making the decision from another's perspective. The experiment validates the hypothesis that participants are sensitive to evidence quality such that strong evidence was far more frequently endorsed than weak evidence. This interpretation supports the conclusion that changing decision perspective allows participants to be sensitive to the available evidence. The manipulation of decision perspective can now be added to other known manipulations of instruction and time pressure that impact the engagement of analytic processing. The positive implication from this is that there is a factor that is external to the individual's reasoning process that can be manipulated to increase analytic processing. A practical application of this finding would be instruct people to think about contentious issues from another perspectives perhaps a third party that is thought to be impartial. By attempting to adopt that impartial perspective individuals should be less biased in their conclusions.

Belief bias is a resilient effect that is challenging to reduce (Evans et al. 1994, Newstead et al. 1992). The regression analyses of belief indicated that the portion of the variance accounted for by belief was greater from the personal perspective than from the researcher's perspective.

While manipulating perspective did decrease belief bias measures of individual differences were included to explore their effects as well.

Reliance on evidence quality from both perspectives was related to estimated IQ among the 77 participants in this analysis. This was strong evidence for IQ indexing the algorithmic level (Stanovich, 2009). The reflective mind was one of the chief areas of investigation and it was here that the AOT results provide some compelling findings. A preference for active open-minded thinking was positively associated with using evidence quality in decisions from the researcher's perspective. That same active open-minded thinking preference was negatively associated using personal beliefs from the participant's perspective. Of all the individual differences measures, the AOT results most closely resemble the theorized reflective level of Stanovich (2009).

The CRT measures a form of cognitive patience and it was related to evidence quality from both perspectives. By these results, the CRT seems to be more closely related to analytic processing rather than heuristic processing which the test items would predict. The heuristically generated responses in this test are all consistently incorrect. It makes theoretical sense for a measure of cognitive patience to be related to the ability to avoid giving intuitive belief based responses.

The EBI, GDMS and the EUA were all included in an effort to explore additional measures of individuals differences not traditionally included within reasoning research. Specific predictions were not made for the expected relationships of these measures to the reasoning outcomes. Of the five subscales of the EBI the simple knowledge and omniscient authority subscales had negative relationships with evidence strength from both perspectives. This is interpreted as individuals who most strongly agreed that "Too many theories just complicate

things." and "People shouldn't question authority." did not rely on evidence quality conversely, those individuals who did not endorse a belief in the simplicity of knowledge had higher evidence strength beta weights. The subscales themselves were consistently inter-correlated with each other but given that most correlated with AOT it is not clear that they were capturing anything beyond that of the AOT.

The GDMS had better performance than the EBI in that more of its subscales had statistical relationships with reasoning performance. An intuitive decision making style was negatively associated with evidence strength beta weights from the participant's perspective. There was a positive relationship between dependent decision-making style and belief beta weights from the researcher's perspective. Those individuals who indicated they preferred to consult other people when making important decisions were likely to use personal beliefs when making the decision on behalf of the researcher and from their own perspective. The avoidant subscale (i.e. I generally make important decisions at the last minute.) was positively related to belief beta weights from the researcher's perspective. Participants who avoid devoting time to the decision making process are not likely to be examining the evidence too closely when they finally make that decision. This interpretation is apt for the spontaneous subscale as well where it was negatively related to evidence strength beta weights from both perspectives. With the exception of the spontaneous subscale none of the significant findings were consistent across perspectives. As in the case of the EBI most of the GDMS subscales correlated with AOT so it is not clear that these subscales are accounting for anything beyond that of the AOT.

The EUA performed in an unexpected manner. The total number of each type of responses was correlated with each beta weight and the evaluative answers were negatively correlated with the use of evidence strength, which is completely counterintuitive. Those

individuals who are evaluative in their epistemological views are exactly the group to be expected to use the evidence in the reasoning problems. On the contrary, the total number of multiplist responses would be expected to correlate with the use of belief but it was positively related to the use of evidence strength from the researcher's perspective. Of the measures included here the AOT, CRT and the Shipley institute of living scale were the most meaningful in their results. The GDMS, EUA and EBI were the least successful of those included measures.

In future investigations of Dual Process theory perspective manipulations combined with individual differences measures like the AOT and CRT may allow for more fine-grained distinctions between the algorithmic mind and the reflective mind. With greater understanding we can investigate if this reflective mind is changeable. In the current study it was treated as if it were a global outlook or tendency. It would be useful to know if individuals would change their perceptions of themselves if certain thinking styles were rewarded in an experimental setting. The use of scientific reasoning problem stimuli such as these can be expanded to investigate how individuals interpret popular media based science reporting. These stimuli only included relevant information for the requested evaluation it would be of interest to explore how participants filter out irrelevant information. Another extension would be to examine how individuals weight findings in light of research sponsor identities given the increasing number of business and science partnerships.

Earlier it was stated that the purpose of the current study was to investigate variables that underlie propensity to override a heuristic TASS response and intervene with analytic response in a scientific evaluation task. It was found that the manipulation of decision perspective effected reasoning performance in both the analyses of belief bias and for evidence quality. The results demonstrate that this simple manipulation of perspective increases the use of analytic processes

in reasoning performance. Individual differences in thinking style measured by the AOT and CRT were related to using personal beliefs and evidence quality in making conclusions in a scientific reasoning task. These measures of thinking disposition are suggested as indices of the reflective mind (Stanovich, 2009).

5. Conclusions

This study contributed to the investigation of the dual process theory of reasoning and the Reflective mind as proposed by Stanovich (2009). The use of a scientific evaluation task allowed for a novel manipulation of evidence quality and decision perspective. These findings add to the growing body of literature on decision perspective (Beatty & Thompson, 2007; Thompson Evans, & Handley, 2005; Greenhoot, Semb, Colombo, & Schreiber, 2004). The extensive inclusion of individual difference measures provided confirmation that the popular actively open-minded thinking scale is theoretically interesting when compared to a diverse selection of alterative measures.

6. References

Beatty, E. L & Thompson, V.A. (2007). Influences on conclusion endorsement: Perspective and belief bias. Paper presented at: CSBBCS 2007. The 17th Annual Meeting of the Canadian Society for Brain, Behaviour and Cognitive Science (BBCS), Victoria, BC. Abstracts of the 2007 CSBBCS Annual Meeting Résumés du congrès annuel 2007 de la SCCCSC. *Canadian Journal of Experimental Psychology, 61,* 363.

Bowers, T.L. & Pantle, M.L. (1985). Shipley institute for living scale and the kaufman brief intelligence test as screening instruments for intelligence. *Assessment, 5*, 187-195.

De Neys, W., Vartanian, O. & Goel, V. (2008). Smarter than we think: When our brains detect that we are biased. *Psychological Science, 19,* 483-489.

Evans, J.St.B.T. (2009). How many dual process theories do we need; One, two or many? In J.St.B.T. Evans & Frankish (Eds) (in press). I*n two minds: Dual processes and beyond.* Oxford: Oxford University Press.

Evans, J.St.B.T., Newstead, S.E., Allen, J.L. & Pollard, P. (1994). Debiasing by instruction: The case of belief bias. *European Journal of Cognitive Psychology. 6,* 263-285.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19,* 25-42.

Greenhoot, A.F., Semb, G., Colombo, J., & Schreiber, T. (2004). Prior beliefs and methodological concepts in scientific reasoning. *Applied Cognitive Psychology, 18*, 203-221.

Jarmasz, J. & Hollands, J.G. (2009). Confidence intervals in repeated-measures designs: The number of observations principle. *Canadian Journal of Experimental Psychology, 63,* 124-138.

Klaczynski, P. A., & Gordon, D. H. (1996). Self-serving influences on adolescent's evaluations of belief-relevant evidence. *Journal of Experimental Child Psychology, 62,* 317–339.

Klaczynski, P.A., & Lavallee, K.L. (2005). Domain-specific identity, epistemic regulation, and intellectual ability as predictors of belief-based reasoning: A dual-process perspective. *Journal of Experimental Child Psychology, 92*. 1-24.

Klaczynski, P.A., & Narasimham, G. (1998). Development of scientific reasoning biases: Cognitive versus ego-protective explanations. *Developmental Psychology, 34*, 175-187.

Klaczynski, P. A., & Robinson, B. (2000). Personal theories, intellectual ability, and epistemological beliefs: Adult age differences in everyday reasoning tasks. *Psychology and Aging, 15,* 400-416.

Kuhn, D. (1991). *The skills of argument*. Cambridge: Cambridge University Press.

Kuhn, D., Cheney, R., & Weinstock, M. (2000). The development of epistemological understanding. *Cognitive Development, 15,* 309-328.

Newstead, S.E., Pollard, P., Evans, J.St.B.T., & Allen, J.L. (1992). The source of belief bias effects in syllogistic reasoning. *Cognition, 4,* 257-284.

Risen, J.L., & Gilovich, T. (2007). Another look at why people are reluctant to exchange lottery tickets. *Journal of Personality and Social Psychology, 93,* 12-22.

Sá, W. C., West, R. F., & Stanovich, K. E. (1999). The domain specificity and generality of belief bias: Searching for a generalizable critical thinking skill. *Journal of Educational Psychology, 91*, 497-510.

Sá, W., Kelley, C., Ho, C., & Stanovich, K. E. (2005). Thinking about personal theories: individual differences in the coordination of theory and evidence. *Personality and Individual Differences, 38,* 1149-1161.

Schraw, G. Dunkle, M.E. & Bendixen, L.D. (1995). Cognitive processes in well-defined and ill-defined problem solving. *Applied Cognitive Psychology, 9,* 523-538.

Scott, S. G. & Bruce, R.A. (1995). Decision-making style: The development and assessment of a new measure. *Educational and Psychological Measurement, 55,* 818-831.

Stanovich, K. E. (2009). Distinguishing the reflective, algorithmic, and autonomous minds: Is it time for a tri-process theory? In J. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 55-88). Oxford: Oxford University Press.

Stanovich, K. E., & West, R. F. (2007). Natural myside bias is independent of cognitive ability. *Thinking & Reasoning, 13,* 225-247.

Stanovich, K. E., & West, R. F. (1997). Reasoning independently of prior belief and individual differences in actively open-minded thinking: the argument evaluation task. *Journal of Educational Psychology, 89*, 342-357.

Stanovich, K. E., & West, R. F. (1998). Individual differences in rational thought. *Journal of Experimental Psychology: General, 127*, 161–179.

Torrens, D., Thompson, V. A., & Cramer, K. (1999). Individual differences and the belief-bias effect: Mental models, logical necessity, and abstract reasoning. *Thinking and Reasoning, 5,* 1-28.

Thompson, V. A. (2009). Dual process theories: A metacognitive perspective. In J. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 171-196). Oxford: Oxford University Press.

Thompson, V. A., Evans, J. St. B. T., & Handley, S. J. (2005). Persuading and dissuading by conditional argument. *Journal of Memory and Language, 53*, 238-257.

Weiss, J.L, Schell, R.E. (1991). Estimating WAIS–R IQ from the shipley institute of living scale: A replication. *Journal of Clinical Psychology, 47*, 558-562.

West, R. F., Toplak, M. E., & Stanovich, K. E. (2008). Heuristics and biases as measures of critical thinking: Associations with cognitive ability and thinking dispositions. *Journal of Educational Psychology*, *100*, 930-941.

Zachary, R. A. (1991). *Manual of the shipley institute of living scale*. Los Angeles: Western Psychological Services.

Zachary, R.A, Crumpton, E., Spiegal, D.E. (1985). Estimating WAIS-RIQ from the shipley institute of living scales. *Journal of Clinical Psychology, 41*, 532-540.